Report: Humanize Data Science Curriculum

Di Wu

Spring 2023

Contents

1 Background

1.1 Heart of Education

> "Placing students at the heart of their education fosters a rich environment for learning and provides the opportunity for students to cultivate attention, deepen their understanding of their studies, engender richer relationships with themselves and others, and stimulate profound inquiries into the nature of themselves and the world around them."
>
> – Daniel Barbezat and Mirabai Bush, *Contemplative Practices in Higher Education (2014)*

We created the Master of Science in Data Science program at CU Boulder with the knowledge that our students are highly motivated by their specific needs: acquiring the necessary skills, techniques, transcript, and degree to secure a well-paid job. Our students typically have some work experience before joining the program, and view it as a checklist of tasks to complete. Our program has been tremendously successful in achieving its objectives: we impart fundamental theories as well as cutting-edge techniques, encourage hands-on projects as well as teamwork, and value curiosity and self-exploration.

While Data Science is the primary focus of our program, I have come to realize that our students are equally important. In addition to gaining knowledge in Data Science, they must also learn from their classmates, faculty, staff, industry practitioners, and themselves. While Data Science is an objective field, people and human interactions are subjective. We can prepare our students to be competitive candidates for job positions, good employees for their jobs, and successful data scientists in the field. However, we can also address the essence of education and train our students to be great people who are curious, passionate, and possess the ability to explore dynamic environments and adapt to the future.

Therefore, I propose the adoption of a Humanized Data Science curriculum that serves both of these objectives more effectively.

2 Overview

The overview of the curriculum is as follows in Table 1

Subject	Tools Kit	Pedagogical Concept	Accomplished?
All	Design Flow	Backwards Design	Yes
All	Center of the Curriculum	Subject-Oriented Teaching	Yes
All	Teaching Philosophy	Establishing Expectation	Yes
All	Open Education Resource	Inclusive Design	Yes
All	Flipped Classroom	Student Contribution	Yes
All	Cheat-sheet Competition	Student Contribution	Yes
All	Reflective Writing	Slow Pedagogy, Metacognition	Yes
All	Dialectic Exchange	Subject-Oriented Teaching	Not yet
All	Meal Interview	Student/Industry Voice	Yes
All	Course project	Subject-Oriented Teaching	Yes
Course	Survey in Exams	Student Voice	Yes
Students	Meditation	Metacognition	Yes

Table 1: Summary of Report

3 Design Flow

To achieve this goal, I followed the Backward Design Approach and began with the question of what students should aim to become upon graduation. The answer is straightforward: a competent team member with a desirable *personality* and a *skill set* that aligns with their domain-specific position and *learning ability*. Accordingly, assessments, activities, and course contents will be tailored to this desired outcome.

The execution of the design flow is as follows in Table 2.

Outcome	Assessment	Grade Percentage	Activity
Course material	Individual	9	Theory
Application	Individual	15	Application
Learning	Class	6	Learning Competition
Application, Learning	Individual	10	Mini Projects
Personality, Learning	Team	10	Team Discussion
All	Team	5	Project Debut
All	Team	5	Project Checkpoint
All	Team	10	Final Presentation
All	Team	10	Final Report
Learning	Individual	0	Case Studies
Course Material, Learning	Individual	20	Exams
All	Team	3 (extra credit)	Investment Game

Table 2: Backward Design

4   Center of the Curriculum

The curriculum is designed using the subject as the center. Based on the outcome of the program, the subjects are: personality, course material, and learning ability.

4.1   Personality

Students are expected to acquire the ability to interact effectively with others, which includes communication, social, teamwork, and conflict management skills that are essential for their career. These skills are not only necessary for interacting with colleagues but also with clients and customers. Moreover, they should know how to present themselves in public and collaborate with their colleagues to achieve team goals. They should also be equipped to handle emergencies or conflicts among interest parties. Therefore, personality is the most crucial trait, even more than the other two subjects. We aim to prioritize this aspect explicitly and depict the depth and breadth of how to improve personality in Figure 1.
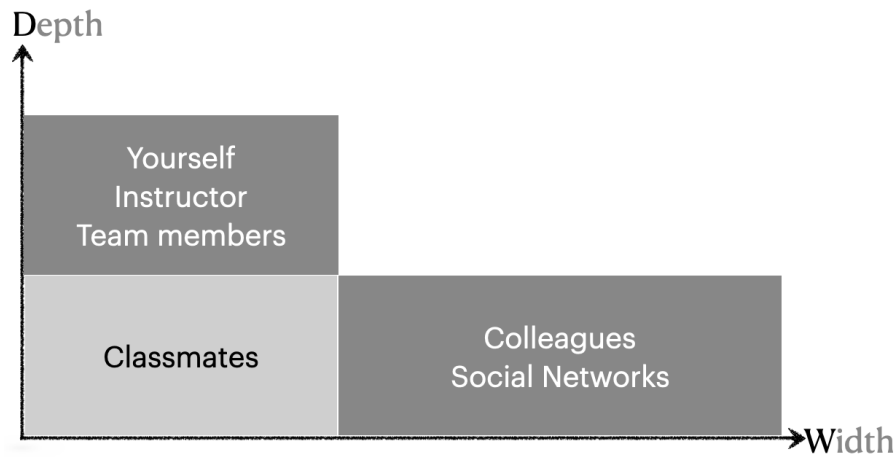
Figure 1: How to Improve Personality

4.2 Learning Ability

Students are expected to excel in learning as Data Science is a field that evolves rapidly and has a highly dynamic environment. While we aim to teach them all the necessary knowledge, it's impossible to teach them new concepts that do not exist currently. Therefore, they must be adaptable and able to learn independently. Improving their learning ability is crucial to their success. This includes not only the acquisition of knowledge but also the development of effective learning strategies. Figure 2 outlines the scope and depth of our approach.
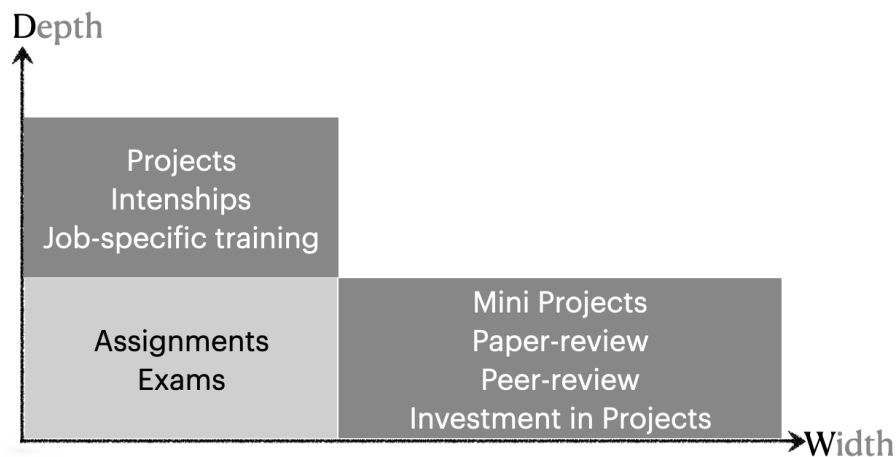

Figure 2: How to Improve Learning Skills

4.3 Course

The course material is an essential component of the curriculum. However, I view it as a means to train students to be good people and good learners, rather than an end in itself. In other words, while the course material is important, it is the least important of the three

subjects. The course content covers both theoretical and practical aspects of data science. I aim to expose students to a wide range of topics, delve deeply into some of them, and leave others for them to explore on their own. Figure 3 illustrates the depth and breadth of how we can enhance students' understanding of the course material.
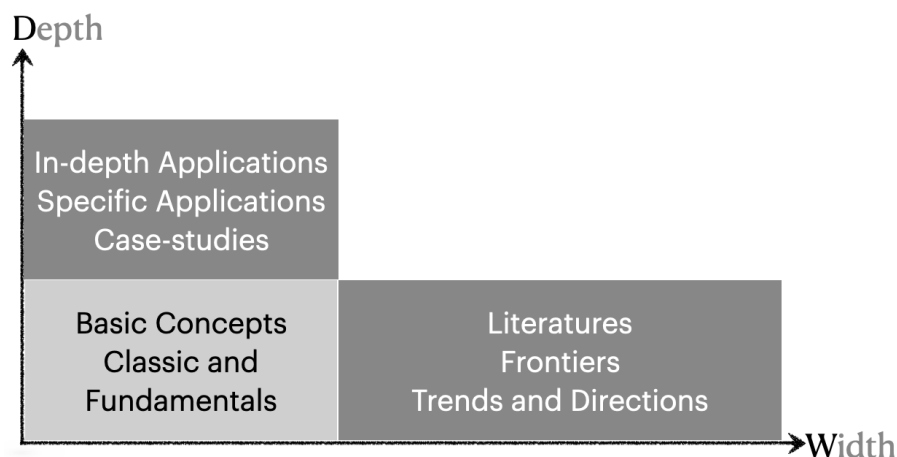

Figure 3: How to learn the Course

4.4   Report

I am able to successfully coordinate these three centers of the curriculum. The activities and their weightage towards the final grade are detailed in Table 2. The detailed explanation is included in my teaching philosophy presentation slides. Please see Appendix A.

5   Teaching Philosophy

To establish the expectations of the course, I'd like to develop a 30-min presentation addressing my teaching philosophy in the first class. With this presentation, students will know exactly what I expect from them and can plan accordingly.

5.1   Report

The presentation proved to be an effective introduction to my courses, providing students with a detailed understanding of my approach to teaching, as well as insights into my personal interests, such as my fondness for World of Warcraft. During the presentation, I emphasized my definition of learning and my expectations for students to be diligent learners. The presentation slides is included in Appendix A.

Looking ahead, I plan to create a video recording of the presentation and make it available on Canvas as a brief introduction to the course. I anticipate completing this task in summer 2023.

6 OER

To remove the barrier of the cost of textbooks and promote inclusivity, I intend to create an Open Educational Resource for my Data Mining course and implement it in the Spring 2023 semester.

6.1 Report

With the professional and patient support of OER Leads, I have successfully completed the OER project. The course material, entitled "Data Mining with Python," is available for free and can be accessed via CU Experts (Link). Here is the brief introduction of the material:

Why we need this book Data is everywhere and it's growing at an unprecedented rate. But making sense of all that data is a challenge. Data Mining is the process of discovering patterns and knowledge from large data sets. This book focuses on the hands-on approach to learn Data Mining. This book is designed to give you an understanding of Data Mining concepts in an applicable way. The tutorials in this book will help you to gain practical skills to implement Data Mining techniques in your work. Whether you are a student, a data scientist or a business analyst, this book is a must-read for you.

How to use this book This book is served as a complementary to a theoretical Data Mining course. We intend to keep the introductions brief and simple, and concentrate in detailed tutorials.The book are divided into two parts:

- Part 1 covers the preparation of data, or Data Wrangling. Part 1 has chapter 1 to chapter 5.

- Part 2 covers the analysis of data, or Data Analysis. Part 2 has chapter 6 to chapter 10.

Please find the associated tutorials in each folder. When you run some .ipynb files, if applicable, please make sure the data path is updated in your local/cloud environment. It will be ideal if you not only run the tutorial, but also change parameters and observe the difference. That is the best way to learn. You can download the NotebookTutorials.zip and run the .ipynb files on your device or via cloud services.

7 Flipped Classroom

To foster a subject-centered community, I intend to implement several student-led lectures in the form of a flipped classroom activity. Students will be presented with a list of topics from which they can select, design a 15-minute lecture, and provide a 15-minute tutorial. The audience will provide feedback, enabling students to refine their skills and improve their performance.

7.1 Report

I encountered some challenges in my Data Mining course due to its large class size (50+) and limited time. However, I successfully implemented a student-led lecture activity in my Data Structure course, which proved to be very effective. In each class session, students were given 7 to 10 minutes to introduce a topic of their choice, after which I continued the lecture based on their presentation. This activity not only engaged the students but also demonstrated that the classroom was a collaborative space. It also encouraged them to learn how to find resources, deliver knowledge, and allocate their time wisely. However, to ensure the success of this activity, it is crucial to provide clear and detailed instructions to the students; otherwise, they may feel confused and lack direction.

The assignments are included in Appendix B.

8 Cheat-sheet Competition

Summarizing a subject is an important aspect of the learning experience and serves as a valid assessment for achieving learning outcomes. As such, I will require students to prepare cheat-sheets for their exams, and hold a competition for the best cheat-sheet. The winner will be selected through a vote among both the students and the instructor. This competition will motivate students to strive for excellence and improve their understanding of the course material. Moreover, the peer-review process will allow students to learn from each other's strengths and weaknesses, which will enhance their knowledge in all three subjects.

8.1 Report

As this activity proved to be highly successful during the fall 2022 semester, I decided to break it down into several smaller activities for the spring 2023 semester. For weekly assignments, there will be a section called "Learning in Cohort," which requires students to find and share the best resources available for a given topic in discussion board. They must also provide reasons for choosing their selected resources. The most popular resource, as determined by the number of likes, will earn its author a small reward, such as being exempt from the next assignment.

9 Reflective Writing

I am thinking that, maybe we can add another dimension to our curriculum: reflection. Reflective writing is an effective approach to help students to better understand the curriculum, better understand the learning experience, and better understand themselves and their surroundings.

9.1 Reflection

"For reflection to be considered critical it must have as its explicit focus that uncovering, and challenging of power dynamics that frame our decisions and actions. It also attempts to challenge hegemonic assumptions; those assumptions

we embrace as being in our best interests when in fact they are working against us. This is what makes critical reflection truly critical."

– Stephan Brookfield, *So Exactly What is Critical about Critical Reflection?* (2016)

For lectures, assignments, and mini projects, reflective writing can help students to look through their learning journey and ask some specific questions:

- How did you understand this lecture?
- If you had difficulties in learning this concept, what are they? How did you overcome them?
- What was your experience working with your teammates? Did you have hard moments for coordinating with each other? How did you reach agreements and got the work done?
- What is your take-away of today's lecture? Why is it? If you were the instructor, how would you make a question and test it?

Reflective writing can help students to slow down, take a pause, and think about some big questions:

- Why do you study Data Science? Why do you want a job in Data Science?
- What do you want after successfully graduating and taking the job?
- How to stay current? How will you apply the learning techniques for new techniques in the future?

Reflective techniques we can have for this task are Loop Writing and Collaborative Learning Questions for individual thinking, Dialectical Notebooks and Writing in the Zone in a group for team thinking, and Process Writing for a deeper and complete thinking over all.

9.2   Dialectical Exchange – An Example of Reflective Writing

Most of our courses require a team course project. In the past, students will receive an oral feedback from their colleagues, and a written feedback from instructors. I would like to enhance the feedback part of the course project, to make it more meaningful and useful, as an important activity in our courses.

Besides the normal requirement of project presentations and project reports, a project Reflective Writing will be required. It will be consisted by four sections.

Part 1: Self Reflection on Course Project   Students on Team A should reflect their learning and working experience on this course project at first. They can address the motivations, the initial ideas, the design, the difficulties, and their strategy to deal with these difficulties, and their tips and lessons for other teams. I will prepare a detailed and specific instruction so students will write reasonable amount of reflection.

Part 2: Peer Comments on Part 1   Part 1 will be assigned to another team, say Team B for their review. The rubric will be provided to both teams. Team B should address whether or not they have similar difficulties, if yes, how did they deal with them, and if not, how did they manage to avoid them, and what are their thoughts about the tips and lessons shared by Team A, what are their suggestions for Team A, etc.

Part 3: Peer Comments on Part 1 and 2 as a Conversation   Both Part 1 and Part 2 will be assigned to Team C for review. Team C will read the reflections of Team A and B, and summarize and comment on the conversation. They can follow the dialog, and contribute their opinions as well as evaluate the effectiveness of the conversation.

Part 4: Self Reflection on Part 1/2/3 as a Dialog   All three parts will be passed back to Team A. They need to read them carefully, and use them to polish and revise their part 1. They need to find out whether or not they delivered a clear and complete information to their audience, and whether or not they get the help they need. The revised reflection together will the part 1 - 3 will be submitted as part of their final report.

9.3   Report

The use of reflective writing has been successful in achieving its intended purpose. Various activities in the course, including team discussions, theoretical portions of assignments, and summary and reflection components of exams, involve reflective writing. At the outset, students often struggle with this type of thinking and writing. However, with regular and consistent practice, I believe they have come to appreciate the value of reflective writing. My hope is that they will continue to develop this habit even after graduating.

Some example questions I used for team discussion as:

- What are the *three* things you learned in today's class? What are the *two* things DIFFERENT than your understanding before today's class? What *one* topic you'd like to in Exam 2?

- What are the *three* similarities between classification and regression? What are the *two* differences between them? Explain *one* similarity or difference in detail with examples.

- From four categories of clustering methods we learned, choose *three* for comparison with examples. How the *two* partitioning method, K-Means and K-Medoids differ? Explain *one* real life scenario that we can apply DBSCAN.

- What is your *one* sentence summary of this presentation (by a team)? Write *one* aspect that you like most of their presentation. Provide *one* suggestion that you think the team should improve.

Unfortunately, the Dialectical Exchange activity has not yet been implemented due to the large class size and limited time. I will look for opportunities to introduce this activity in the future if conditions permit.

10 Meal Interview

In order to keep current, I want to create and maintain a routine which is a stable, open, and concentrated channel of getting inputs. Here I learned from Dr. Michael Wesch in his TEDxMHK talk in which Dr. Wesch talked about that he conducts lunch interview with students every weekday. I plan to have weekly/biweekly meetings with students, industry practitioner,and academic colleagues. During the meetings, we will not have small talks but directly get to the business. To encourage open and frank communication, I'll respect their privacy and keep them anonymous.

10.1 Report

I found conducting interviews to be an extremely valuable experience. Without sitting down with my students and spending an hour talking with them, I wouldn't have been able to gain such a deep understanding of their true feelings. These interviews also helped me discover some excellent dining locations on campus, such as C4C. Furthermore, the interviews, even though they took place in a restaurant, set up an expectation that we would be discussing serious topics. Colleagues, particularly practitioners in the industry, went above and beyond my expectations.

Some takeaways from the interviews are:

- I know many students are beginners, and you are trying to make everyone comfortable. However, the practices and assignments mandatory required are too easy and too boring for me. I had industry experience and I am way advanced than others, but I still have to do the assignments. I felt wasting my time.

- I like how you designed the course. It is unique and I love to learn by doing. I wish we could do in-class tutorial for every topic we learn.

- You should be more assertive to students. We are from a culture that professors won't smile at you.

- You did a wonderful job in explaining theoretical background of the topics, but I want to see more real life examples and intensive coding.

- The optional mini projects are great, but I just don't have time to do them. Please make them mandatory.

- For the lectures, I really like yours. But, you should look at [another Professor]'s assignments. They are very challenging, but I feel very good after I get them done.

- I understand team project is good for learning. I just don't like team work, and I don't want to push team members to do the work. Either I don't care, or I have to do everything. I prefer do the project by my own.

- Your slides, videos, and other course materials are so complete (so sometimes I feel I will still be fine if I skip some classes).

- We hire talent, not a tool. We actually do not expect the student knows everything. However, we want to hire someone who can learn quickly (in a team).

- Our company is in a bad position, and it will remain falling next year (2023). Students should be prepared to compete with laid-off employees.

- A lot of cases, what you have done is not that important, but what you can deliver is.

- My department [English and Communication] actually cares about data, and data analysis. We learned and use Machine Learning and AI for research and teaching too.

- Everyone is busy. If you have a problem, ask around, but don't just wait for answers. Do some research quickly by yourself.

Based on these (and many other) comments, I made adjustments to my courses:

- I modified assignments, and made some questions optional but more challenging. By modifying the assignments, you have given students more choices and made the class more accessible to students with different skill levels and backgrounds. The optional challenging questions provide an opportunity for students who want to push themselves, while the mandatory questions ensure that students receive the necessary knowledge.

- I added many applicable, case-study style, hands-on tutorials for the content this semester. It is a great way to help students apply what they've learned and make the content more engaging. Students valued them, and listed them as the favorite.

- For 20 mini projects (almost 2 mini projects per week), I set up 10 as the threshold to get the full credit. Students don't have to do all, but they cannot drop all. This, again, provides flexibility. By setting a threshold, I encouraging students to complete a reasonable amount of work while still allowing them to choose which projects they want to focus on. The other thing I learned, is that students want to have all projects unlocked at the same time, so they can plan ahead, even these projects are due in different weeks.

- I delivered the comments from industry practitioners. It helps prepare students for real-world situations and gives them a better understanding of what is expected of them in their future careers.

- I added some in-person only activities, which are also a good way to engage students who might otherwise prefer to attend classes online.

- I weighted a lot in their presentations, and this emphasis on communication skills and presentation ability is a valuable lesson for students. It's important for them to understand that being able to communicate their ideas effectively is just as crucial as being able to code..

11   Course Project

Course projects have been the core component of all my courses. I would like to improve the course projects so all three subjects can be addressed.

11.1 Report

Course projects are essential for both their contribution to the final grade and their role in helping students achieve all the outcomes outlined in the course. This semester, while there are no constraints on the scope of course projects, I have provided detailed instructions. To facilitate student learning, I have started to candidly point out their errors and imperfections, even though I know this may be frustrating for them. It's important to recognize that happiness does not equate to wellness. My hope is that students can learn from their mistakes and imperfections and use that knowledge to improve their performance in future projects, ultimately preparing them for success in the real world.

12 Survey in Exams

I include a mandatory survey in exams and the survey worth 5% of the grade. In this way, I can get feedback from students, and make some changes in the course so students know their opinions matter.

12.1 Report

The survey has become a regular part of my exams. This semester, I made the survey anonymous to encourage students to provide more candid feedback. Similar to the meal interviews, the survey has been incredibly beneficial in helping me understand my students' perspectives. Most importantly, when students see that I make immediate adjustments to the course design based on their feedback, they feel acknowledged and are more willing to collaborate with me to enhance the course. Additionally, each semester brings a different cohort of students, and while some activities may be well-received by the current cohort, I must remain open to making changes in the future based on feedback from subsequent cohorts.

The questions I asked in the survey are:

- How do you feel about this exam? How shall we improve it for exam 2?

- How do you feel about our homework so far? What should be improved for the rest of the homework?

- How do you feel about our class so far? What should be improved for the next half of the course?

- How do you feel about our mini projects so far? What should be improved for the next semester? For the 17 projects, shall we make 10 rather than 7 as required?

- How do you feel about our class so far? What should be improved for the next semester?

- What is your most favorite activity in this class? What is the least?

13 Meditation

I used to play some random music (piano, cello, jazz, sci-fi, etc.) before the lecture starts. I plan to create a list of music that helps students put their busy life down and be ready to embrace our meetings.

13.1 Report

It is understandable that not all students may have a liking for piano music, and some may even dislike it. Hence, instead of playing piano music before lectures, I opt to play recent news and connect them to the course material. By doing so, I encourage students to question how the news relates to the lecture, which in turn promotes engagement during my lectures.

14 Conclusion

This project is a significant one with numerous details that need to be filled, tested, and modified. It is unrealistic to expect that all the ideas will come to fruition in one semester, one year, or even several years. This project lays out a path towards my ideal curriculum, the Humanized DS Curriculum. I am grateful that most of the tasks were implemented and achieved success in Fall 2022 and Spring 2023. I will continue to experiment with these ideas and strive to make them even better.

15 Acknowledgement

I would like to use this section to acknowledge the people who have supported me throughout this project:

- ASSETT 2022-2023 Program: I feel fortunate to have been selected for this one-year program. I have learned a multitude of tools and methods that have greatly helped me in my pedagogical practice. The biweekly meetings in Fall 2022 were always the most anticipated time for me, as I got to meet with colleagues who share my passion for education and care for our students. Amanda Seal McAndrew, you did a wonderful job organizing the meetings, inviting experts, and handling all kinds of additional questions. You are truly a gem in the program. Kirk Ambrose, Mark Arnett, Lauren Collins, Christy Fillman, Emily Greenwood, Jillian Porter, Tracy Quan, and Nicole Stob, thank you for sharing your experiences, ideas, challenges, and feelings. I enjoyed our discussions, and I am fascinated by how each of us chose a different path for the same icebreaker questions.

- Computer Science Program: As a new member of the CU Boulder system, I encountered many problems in setting up my fund, applying for a P-card, or using the P-card for purchases, among other things. The CS team, especially Stephanie Morris, is the most supportive team anyone could hope for. They have always helped me with patience, information, and detailed explanations.

- OER Leads: Developing this OER has been a wonderful learning experience for me. I have gained a lot of knowledge about copyrights, licenses, formatting, and how to properly use other OER materials. Merinda McLure and Caroline Sinkinson, I appreciate your knowledge, support, and patience.

- My colleagues and students in the MSDS program: I am grateful to be a member of this program, surrounded by so many people from all over the world. I appreciate each and every student, colleague, and friend in the industry. Thank you for your time, authentic suggestions, and some challenging criticisms. I have enjoyed every conversation we had over the past year, every comment you gave me in various ways, and every moment we nodded and smiled. Thank you.

A Teaching Philosophy

Teaching Philosophy
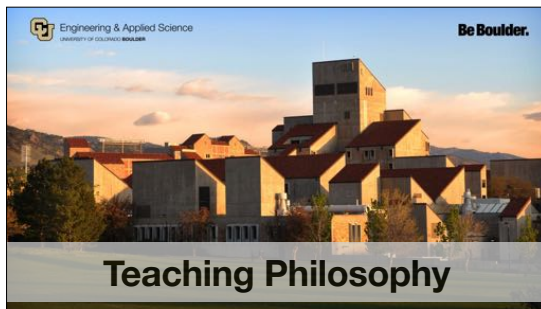
Learn - What - Subject Matter

- How to think
- How to talk
- How to explain
- How to apply
- How to extend
- How to shine

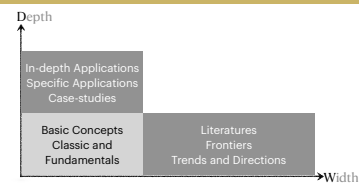University of Colorado **Boulder** — 2 — Di Wu

Learn - What - Beyond

- Deal with learning
- Deal with choices
- Deal with difficulty
- Deal with pressure
- Deal with people
- Deal with yourself

University of Colorado **Boulder** — 3 — Di Wu

Subject

Depth

In-depth Applications
Specific Applications
Case-studies

Basic Concepts
Classic and
Fundamentals

Literatures
Frontiers
Trends and Directions

Width

University of Colorado **Boulder** — 4 — Di Wu

Learn - Who

- You are the host
- Be you
- Rely on you
- Take the responsibility
- Accept the consequence
- No excuses

University of Colorado **Boulder** — 5 — Di Wu
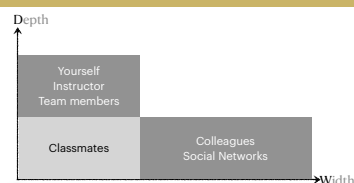
Learn - Whom

- Learn from yourself
- Learn from the instructor
- Learn from colleagues
- Learn from competitors
- Learn from strangers
- Learn from anyone

University of Colorado **Boulder** — 6 — Di Wu

People

Depth

Yourself
Instructor
Team members

Classmates

Colleagues
Social Networks

Width

University of Colorado **Boulder** — 7 — Di Wu

Learn - How - Basic

- Be Active
- Use the time wisely
- Meet the expectation
- Do your part
- Take something away

University of Colorado **Boulder** — 8 — Di Wu

Learn - How - Advanced

- Be proactive
- Organize time conflicts
- Go beyond the expectation
- Lead / care others
- Contribute something

University of Colorado **Boulder** — 9 — Di Wu

Teach - What

- Subject Matter:
  - Width/Broad V.S. Basic/Specific
  - Depth/Specific V.S. Basic/Fundamental
  - Theoretic/Research V.S. Applicable/Practical
- Beyond
  - How various way of teaching
  - How various way of learning
  - How various way of failure / success

University of Colorado **Boulder** — 10 — Di Wu

Teach - Approach

Do

Projects

Lectures
Assignments
Mini Projects

Presentations
Peer-review

Watch
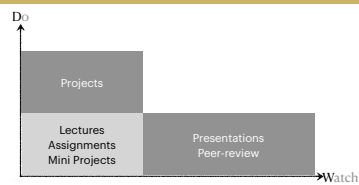
University of Colorado **Boulder** — 11 — Di Wu

Teach - How

- Set up an environment of learning
- Set up various level of requirements
- Set up differentiated activities

Activity	Required	Optional	Observing	Doing
Lectures	X		X	
Graded Assignments	X			X
Ungraded Assignments		X		X
Exams	X			X
Discussions	X	X	X	X
Projects	X	X	X	X

DO MORE. *You can WIN if you want*

University of Colorado **Boulder** — 12 — Di Wu

Teach - Activities

Depth

Projects
Intenships
Job-specific training

Assignments
Exams

Mini Projects
Paper-review
Peer-review
Investment in Projects

Width
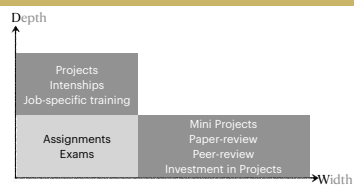
University of Colorado **Boulder** — 13 — Di Wu

Instructor - Di (Dee)

Depth

Education
Data
WoW
Gardening

Di (Dee)

Nature, Culture, Movies, Arts

Width
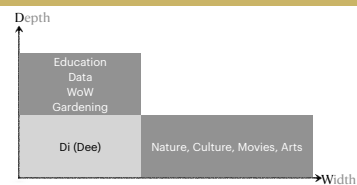
University of Colorado **Boulder** — 14 — Di Wu

Let's Get Started!

B   Flipped Classroom Examples

Project 2: Flipped Classroom (Round 1)

✅ **Published**   ✎ **Edit**   ⋮

Project Background:

This project will be conducted in teams. It will require you to work together and divide and conquer.

Project Assignment:

We have 11 teams. Each group is assigned one type of data structure as below. The data structure with *(Python)* is Python Specific; otherwise, it is a general data structure.

- Round 1 Team 1 (Mar.7): StaticArray, DynamicArray in General
- Round 1 Team 2 (Mar.7): List *(Python)*, Tuple*(Python)*
- Round 1 Team 3 (Mar.9): LinkedList (Singly Linked List, Doubly Linked List, Circular Linked List)
- Round 1 Team 4 (~~Mar.16~~ **Mar.14**): String *(Python),* Stack/Queue

Project Requirements:

- Each group should prepare an Explain-{The Data Structure of Your Team}-in-10-Minutes presentation (you can prepare a video, or you can do a live show). In these 10 mins, you should cover:
    - What is the data structure?
    - What are the **daily-life examples** of this structure (NOT in programs)?
    - What is the **usage** of this structure in Programmings?
    - What are the **operations** of the data structure, and what are their **complexities**?

Project Submission:

You should submit the work **24 hours before** the date that we will cover the data structure.

You can upload the slides or videos to your Google Workspace, and share the link in the submission (please make sure to share it with my **diwu6407@colorado.edu (mailto:diwu6407@colorado.edu)**). I'll play the slides/videos in class and we all will enjoy your tutorial on {The Data Structure of Your Team}.

Points   25

Project 2: Flipped Classroom (Round 2)

✅ Published   ✎ **Edit**   ⋮

Project Background:

This project will be conducted in teams. It will require you to work together and divide and conquer.

Project Assignment:

We have 4 teams. Each group is assigned one type of data structure as below. The data structure with *(Python)* is Python Specific; otherwise, it is a general data structure.

- Round 2 Team 1 (Mar.21): Map(Hash-Table),
- Round 2 Team 2 (Mar.21): Set*(Python),* Dictionary*(Python)*
- Round 2 Team 3 (Apr.4): General Tree, TreeTraverse(Pre-In-Post), Special Trees (Heap Tree, Binary Search Tree, Balanced Binary Tree), BFS/DFS
- Round 2 Team 4 (Apr.11): Graph Representations, Algorithms: BFS/DFS, Shortest-Path, Minimal Spanning Tree

Project Requirements:

- Each group should prepare an Explain-{The Data Structure of Your Team}-in-10-Minutes presentation (you can prepare a video, or you can do a live show). In these 10 mins, you should cover:
  - What is the data structure?
  - What are the **daily-life examples** of this structure (NOT in programs)?
  - What is the **usage** of this structure in Programmings?
  - What are the **operations** of the data structure, and what are their **complexities**?

Project Submission:

You should submit the work **24 hours before** the date that we will cover the data structure.

You can upload the slides or videos to your Google Workspace, and share the link in the submission (please make sure to share it with my **diwu6407@colorado.edu (mailto:diwu6407@colorado.edu)**). I'll play the slides/videos in class and we all will enjoy your tutorial on {The Data Structure of Your Team}.

C   Assignment Example

Assignment 05

✅ **Published**   ✎ **Edit**   ⋮

Objectives

This assignment aims to enhance your understanding of Data Analysis - Classification. By the end of this assignment, you will be able to:

- Differentiate between classification methods, including decision tree, KNN, SVM, etc.
- Apply classification methods and assess results using appropriate evaluation techniques.
- Define hyperparameters, and utilize model tuning methods to identify optimal hyperparameters.

Instructions

- This assignment has three parts: Part A is theoretical, Part B focuses on application, and Part C is learning as a cohort.
- You will prepare multiple PDF files and submit them through Canvas.
- You must work ALONE on this assignment, and while errors are acceptable, you must submit your own work.
- By submitting your answers to this assignment, you acknowledge adherence to the Honor Code Pledge:

    "***On my honor as a University of Colorado Boulder student I have neither given nor received unauthorized assistance***".

- Your submission will be checked for plagiarism, and if any is found, you will receive zero points for the assignment with a warning. If you receive two warnings, you will receive an F grade for this course and be reported to the University.

Part A - Reasoning (30 pts)

Answer the following questions as if you were in an interview. Don't speak too little, but also avoid speaking too much!

1. (10 pts) Discuss the big picture of learning and why we have both supervised and unsupervised learning.

2. (10 pts) Someone claims that the "Nearest Neighbor" method is not a learning method. Do you agree or disagree, and why?
3. (10 pts) Why is Logistic Regression considered a classification method?

*Prepare your answers as a PDF, name it as **your first and last name + A05 + Part A.pdf**, for example, DiWuA05PartA.pdf, and submit it here via Canvas*

Part B - Applications (50 pts)

Load the "titanic" dataset in seaborn (**https://github.com/mwaskom/seaborn-data** ⤷ **(https://github.com/mwaskom/seaborn-data)**) and set "survived" as the dependent variable.

1. (10 pts) Conduct a Decision Tree analysis using scikit-learn library.
2. (10 pts) Conduct a Logistic Regression analysis using scikit-learn library.
3. (10 pts) Conduct a KNN analysis using scikit-learn library.
4. (10 pts) Conduct a SVM analysis using scikit-learn library.
5. (10 pts) Summarize your findings. Explain the results of each method, compare their performance, and discuss the pros and cons of each method.
6. *(optional) Conduct other classification methods for "survived" using scikit-learn library or other relevant libraries.*
7. *(optional) Play with other dependent variables and see how different variables affect the prediction performance.*
8. *(optional) Use the "penguins" dataset in seaborn to predict species and compare the performance of different classification methods.*

You should prepare a tutorial for this part. Your tutorial should include the following:

- Loading the dataset and setting up the dependent variable
- Writing the code for each classification method
- Printing the output and interpreting the results
- Summarizing the findings and discussing the pros and cons of each method

*Print your tutorial as a PDF, name it **your first and last name + A05 + Part B.pdf**, for example, DiWuA05PartB.pdf, and submit it here via Canvas.*

Part C - Learning Resources (20 pts)

Identify **ONE** resource (a paper, blog, a video, a slide, a real-life story, etc) for each item listed below:

1. (10 pts) Ethics in Classification: What it is, why it is important, how we can include it in practice.
2. (10 pts) Imbalanced Dataset: Pick one tutorial that explains the impact of imbalanced dataset in classification, and demonstrates how to deal with it.

Post your resources with a short description to the **discussion board (https://canvas.colorado.edu/courses/89640/discussion_topics/1087480)** , such as:

1. Ethics in Classification:
    1. link:
    2. This resource is [fill in your answer]. I like this resource because [fill in your answer]
2. Imbalanced Dataset:
    1. link:
    2. This resource is [fill in your answer]. I like this resource because [fill in your answer]

*Take a screen shot of your post, convert it to a PDF, name it as **your first and last name + A05 + Part C.pdf**, for example, DiWuA05PartC.pdf, and submit it here via Canvas.*

Bonus: The author of the Top ONE Sharing can be exempt from Assignment06.

Top One Definition: We order the posts by # of likes, if there are multiple posts have same # of likes, then it is ordered by the time it was posted.

Points 100
Submitting a file upload
File Types pdf

Due	For	Available from	Until
Mar 9	Everyone	Feb 28 at 12am	Mar 12 at 11:59pm

+ **Rubric**